

Amendments to the Claims:

Please withdraw claims 1-11, 20-30, and 39-49, leaving claims 12-19, 31-38, and 50-57 pending.

This listing of claims will replace all prior versions, and listings of claims in the application.

Listing of Claims:

1. (Withdrawn) A computer implemented method for characterizing a plurality of biological sequences comprising:
obtaining a plurality of models, wherein each of the models represents a classification of biological sequences with structural or functional similarity; determining fitness of the biological sequences to the models; and automatically classifying the sequences according to the distances to the models.
2. (Withdrawn) The method of Claim 1 wherein the plurality of biological sequences have at least 50 sequences.
3. (Withdrawn) The method of Claim 2 wherein the plurality of biological sequences have at least 100 sequences.
4. (Withdrawn) The method of Claim 3 wherein the plurality of biological sequences have at least 100 sequences.
5. (Withdrawn) The method of Claim 3 wherein the models are Hidden markov models.
6. (Withdrawn) The method of Claim 5 wherein the classification is a family and each model represents a family.
7. (Withdrawn) The method of Claim 6 wherein the sequences are protein sequences.

8. (Withdrawn) The method of Claim 7 wherein the distances are E values.
9. (Withdrawn) The method of Claim 8 wherein the step of automatically determining comprises determining a step of determining a threshold for each of the models.
10. (Withdrawn) The method of Claim 9 wherein the step of determining a threshold comprises performing a curve analysis.
11. (Withdrawn) The method of Claim 10 wherein the step of performing a curve analysis comprises determining a point where the e-value curve drops abruptly or flattens.
12. (Original) A computer implemented method for gene characterization comprising: generating libraries of models using structural relationships of known proteins; inputting a plurality of protein sequences;
comparing the plurality of protein sequences with the models;
automatically establishing criteria for assigning the sequences for each model; and
assigning the sequences to the models based upon the criteria.
13. (Original) The method of Claim 12 wherein the models are hidden markov models.
14. (Original) The method of Claim 12 wherein at least 50 protein sequences are predicted protein sequences.
15. (Original) The method of Claim 14 wherein at least 150 protein sequences are predicted protein sequences.
16. (Original) The method of Claim 15 wherein at least 500 protein sequences are predicted protein sequences.
17. (Original) The method of Claim 12 wherein the step of automatically establishing comprises determining a threshold for each of the models.

18. (Original) The method of Claim 17 wherein the step of determining a threshold comprises performing a curve analysis.

19. (Original) The method of Claim 18 wherein the step of performing a curve analysis comprises determining a point where the e-value curves drops abruptly or flattens.

20. (Withdrawn) A system for gene annotation comprising a processor; and a memory coupled with the processor, the memory storing a plurality of machine instructions that cause the processor to perform logical steps comprising obtaining a plurality of models, wherein each of the models represents a classification of biological sequences with structural or functional similarity; determining fitness of the biological sequences to the models; and automatically classifying the sequences according to the distances to the models.

21. (Withdrawn) The system of Claim 20 wherein the plurality of biological sequences have at least 50 sequences.

22. (Withdrawn) The system of Claim 21 wherein the plurality of biological sequences have at least 100 sequences.

23. (Withdrawn) The system of Claim 22 wherein the plurality of biological sequences have at least 100 sequences.

24. (Withdrawn) The system of Claim 23 wherein the models are Hidden markov models.

25. (Withdrawn) The system of Claim 24 wherein the classification is a family and each model represents a family.

26. (Withdrawn) The system of Claim 25 wherein the sequences are protein sequences.

27. (Withdrawn) The system of Claim 26 wherein the distances are E-values.

28. (Withdrawn) The system of Claim 27 wherein the step of automatically determining comprises determining a step of determining a threshold for each of the models.

29. (Withdrawn) The system of Claim 28 wherein the step of determining a threshold comprises performing a curve analysis.

30. (Withdrawn) The system of Claim 29 wherein the step of performing a curve analysis comprises determining a point where the e value curve drops abruptly or flattens.

31. (Original) A system for gene annotation comprising a processor; and a memory coupled with the processor, the memory storing a plurality of machine instructions that cause the processor to perform logical steps comprising generating libraries of models using structural relationships of known proteins; inputting a plurality of protein sequences; comparing the plurality of protein sequences with the models; automatically establishing criteria for assigning the sequences for each model; and assigning the sequences to the models based upon the criteria.

32. (Original) The system of Claim 31 wherein the models are hidden markov models.

33. (Original) The system of Claim 32 wherein at least 50 protein sequences are predicted protein sequences.

34. (Original) The system of Claim 33 wherein at least 150 protein sequences are predicted protein sequences.

35. (Original) The system of Claim 34 wherein at least 500 protein sequences are predicted protein sequences.

36. (Original) The system of Claim 35 wherein the step of automatically establishing comprises determining a threshold for each of the models.

37. (Original) The system of Claim 36 wherein the step of determining a threshold comprises performing a curve analysis.

38. (Original) The system of Claim 37 wherein the step of performing a curve analysis comprises determining a point where the e-value curves drops abruptly or flattens.

39. (Withdrawn) A computer software product of the invention comprising a computer readable medium having computer-executable instructions for performing the method comprising:

obtaining a plurality of models, wherein each of the models represents a classification of biological sequences with structural or functional similarity;
determining fitness of the biological sequences to the models; and
automatically classifying the sequences according to the distances to the models.

40. (Withdrawn) The product of Claim 39 wherein the plurality of biological sequences have at least 50 sequences.

41. (Withdrawn) The product of Claim 40 wherein the plurality of biological sequences have at least 100 sequences.

42. (Withdrawn) The product of Claim 41 wherein the plurality of biological sequences have at least 100 sequences.

43. (Withdrawn) The product of Claim 42 wherein the models are Hidden markov models.

44. (Withdrawn) The product of Claim 43 wherein the classification is a family and each model represents a family.

45. (Withdrawn) The product of Claim 44 wherein the sequences are protein sequences.

46. (Withdrawn) The product of Claim 45 wherein the distances are E-values.

47. (Withdrawn) The product of Claim 46 wherein the step of automatically determining comprises determining a step of determining a threshold for each of the models.

48. (Withdrawn) The product of Claim 47 wherein the step of determining a threshold comprises performing a curve analysis.

49. (Withdrawn) The product of Claim 48 wherein the step of performing a curve analysis comprises determining a point where the e-value curve drops abruptly or flattens.

50. (Original) A computer software product of the invention comprising a computer readable medium having computer-executable instructions for performing the method comprising:

generating libraries of models using structural relationships of known proteins;
inputting a plurality of protein sequences;
comparing the plurality of protein sequences with the models;
automatically establishing criteria for assigning the sequences for each model; and
assigning the sequences to the models based upon the criteria.

51. (Original) The product of Claim 50 wherein the models are hidden markov models.

52. (Original) The product of Claim 51 wherein at least 50 protein sequences are predicted protein sequences.

53. (Original) The product of Claim 52 wherein at least 150 protein sequences are predicted protein sequences.

54. (Original) The product of Claim 53 wherein at least 500 protein sequences are predicted protein sequences.

55. (Original) The product of Claim 54 wherein the step of automatically establishing comprises determining a threshold for each of the models.

56. (Original) The product of Claim 55 wherein the step of determining a threshold comprises performing a curve analysis.

57. (Original) The product of Claim 56 wherein the step of performing a curve analysis comprises determining a point where the e-value curves drops abruptly or flattens.